

HODGERANK IS THE LIMIT OF PERRON RANK

NGOC MAI TRAN

ABSTRACT. We study the map which takes an elementwise positive matrix to the k -th root of the principal eigenvector of its k -th Hadamard power. We show that as k tends to 0 one recovers the row geometric mean vector and discuss the geometric significance of this convergence. In the context of pairwise comparison ranking, our result states that HodgeRank is the limit of Perron Rank, thereby providing a novel mathematical link between two important pairwise ranking methods.

Let $X = [X_{ij}]$ be a $n \times n$ matrix, $\mathcal{K} := \mathbb{R}_+^{n \times n}$ be the open cone of elementwise positive matrices. For $X \in \mathcal{K}$, let $v(X), \lambda(X)$ denote its unique principal eigenvector-eigenvalue pair in $\mathbb{P}\mathbb{R}^{n-1} \times \mathbb{R}$, $X^{(k)} := [X_{ij}^k]$ be its k^{th} Hadamard power. It is known that if the tropical max-times eigenvector $m(X)$ of X is unique in $\mathbb{P}\mathbb{R}^{n-1}$, then it is the coordinate-wise limit of the sequence $v(X^{(k)})^{1/k}$ as $k \rightarrow \infty$ [1].

$$\lim_{k \rightarrow \infty} v(X^{(k)})^{1/k} = m(X)$$

Our first theorem states that the same sequence converges coordinate-wise to the row geometric mean of X as $k \rightarrow 0$.

Theorem 1. For $X \in \mathcal{K}$, define $h(X)_i = (\prod_{j=1}^n X_{ij})^{1/n}$. Then

$$\lim_{k \rightarrow 0} v(X^{(k)})^{1/k} = h(X)$$

This result plays an important role in the context of pairwise ranking, where X is restricted to the subvariety of *multiplicative comparison* or *symmetrically reciprocal* matrices $\{X \in \mathcal{K} : X_{ij} = 1/X_{ji}\}$ [3, 7, 8, 16]. Here X_{ij} measures the multiplicative preference of i over j . A ranking algorithm takes X as input and returns a score vector $s \in \mathbb{P}\mathbb{R}^{n-1}$ by which the items are ranked. The triple $v(X), m(X), h(X)$ are outputs of three ranking algorithms: Perron Rank [16, 17], Tropical Rank [6, 7] and HodgeRank [2, 12]. Perron Rank plays a fundamental role behind the Analytic Hierarchical Process [16], a ranking procedure extensively applied in decision making. On the other hand, HodgeRank is closely related to many pairwise ranking algorithms in the rank learning literature of computer science [9, 11, 12]. A number of papers have been devoted to their comparisons [4, 6, 8, 12, 15, 16, 19]. However, Theorem 1 is the first to show that HodgeRank can be obtained as the limit of Perron Rank, thus providing a mathematical tool for comparing pairwise ranking methods in computer science to those often used in decision making.

Hadamard powers of a multiplicative comparison matrix arise naturally when one attempts to apply Perron Rank to *additive comparison* matrices, those which arise often in computer science applications [9, 10, 11, 12]. These are skew-symmetric matrices $\wedge_2 \mathbb{R}^n$ where A_{ij} measures the additive preference of i over j . They are in natural bijection with

The author would like to thank Lek-Heng Lim, Nouredine El Karoui and Bernd Sturmfels for helpful comments and advices. Research on this project was supported by the National Science Foundation (DMS-1057064) and (DMS-0968882).

multiplicative comparison matrices via the exponential/log map, however, the bijection is not unique since one is free to choose the base value. Changing the exponential base is equivalent to multiplying the additive matrix A by a positive constant, which can be interpreted as changing the measurement unit. This corresponds to taking Hadamard power of the multiplicative matrix X . Thus, if the input is an additive matrix A and *a priori* no base is preferred, one should consider all the Hadamard powers $[\exp(kA_{ij})]$ when applying Perron Rank $v : X \mapsto v(X)$. This yields the *Perron family* of ranking methods $V_k : X \mapsto v(X^{(k)})^{1/k}$ for any $k \in (0, \infty)$, with HodgeRank and Tropical Rank appear as limiting cases. We shall use the notation $\tilde{V}_k : A \mapsto 1/k \log v([\exp(kA_{ij})])$ to denote its log version.

It can easily be checked that the Perron family, HodgeRank and Tropical Rank are projections onto the set of *strongly transitive* matrices [15] or *consistent* [16] matrices, which are rank one matrices identified with the score vectors in $\mathbb{P}\mathbb{R}^{n-1}$ via $w \leftrightarrow [w_i/w_j]$. Its image in $\wedge_2\mathbb{R}^n$ under the log map is the subspace spanned by the matrices $\sum_{j=1}^n e_{ij} - \sum_{k=1}^n e_{ki}$ for $i = 1, 2, \dots, n$, where e_{ij} is the matrix with 1 in the (i, j) -th entry and 0 else, denoted \mathcal{ST} . These matrices are so named since for all triples i, j, k the relation $X_{ij} = X_{ik} \cdot X_{kj}$ holds, which is natural if one assumes that the comparison matrix is obtained from the score in the error-free. Thus one central question in pairwise ranking is the *score recovery problem*: suppose there is a true score w and one observes X , a perturbed version of the true comparison matrix $[w_i/w_j]$. Which method is ‘best’, in some sense, at recovering w ? By specifying an objective function, such as number of items ranked correctly, one can convert the comparison of methods into an optimization problem. This formulation is commonly found in the broader area of rank learning [5, 9, 10], a rich literature on the study of algorithms for ranking with emphasis on large datasets, predicting the rank of new items and ranking the top items correctly (see, for example, [9] and references therein).

A large class of these optimization problems can be rephrased in $\wedge_2\mathbb{R}^n$ as geometric questions on the images under the projection \tilde{V}_k of level sets of the noise distribution. The second result of our paper, Theorem 2, describes the fibers of this projection. It also contributes towards solving whether there exists an objective function in which Perron Rank for $k \neq 0$ or ∞ is the optimal solution over all possible ranking algorithms. This question arised in the literature since HodgeRank is known to be the ℓ_2 -minimizer and Tropical Rank is known to be a special point on the set of ℓ_∞ -minimizers to \mathcal{ST} [7, 12].

Theorem 2. *Let $\mathbf{1} = (1, \dots, 1)^T$ be the all one vector, $\mathbf{0} = (0, \dots, 0)^T$ the all-zero vector. For all $s \in \mathbb{R}/(1, \dots, 1)$, as a set, the fibers of the map \tilde{V}_k satisfy*

$$\tilde{V}_k^{-1}(s) = [s_i - s_j] + \tilde{V}_k^{-1}(\mathbf{0}).$$

The zero-fiber $\tilde{V}_k^{-1}(\mathbf{0})$ can be decomposed as

$$\tilde{V}_k^{-1}(\mathbf{0}) = \bigoplus_{i=1}^n S_i(k) + \mathbb{R} \cdot \mathbf{1}\mathbf{1}^T,$$

where the $S_i(k)$ for $i = 1, \dots, n$ are orthogonals, with

$$\begin{aligned} S_i(k) &:= \{a_i. \in \mathbb{R}^n : a_{ij} < 0, \sum_{j=1}^n \exp(ka_{ij}) = 1\} \text{ for } k \in (0, \infty), \\ S_i(\infty) &:= \{a_i. \in \mathbb{R}^n : a_{ij} \leq 0, a_{ij} = 0 \text{ for at least one } j = 1, \dots, n\}, \\ S_i(0) &:= \{a_i. \in \mathbb{R}^n : \sum_{j=1}^n a_{ij} = 0\} \end{aligned}$$

Since the fibers $\tilde{V}_k^{-1}(s)$ are just translations of the zero-fiber by $[s_i - s_j]$, Theorem 2 implies that the optimal k^* in the score recovery problem considered is independent of the true additive score s . For large k , the zero-fiber is the real amoeba of the tropical (max-plus) eigenvector map [13]. For $k = 0$, the zero-fiber is the set of matrices decomposing into the all-one matrix plus a matrix whose column sum is zero, which is the zero-fiber of the HodgeRank map. Thus we obtain a crude version of Theorem 1.

In summary, Theorem 1 and its geometric cousin, Theorem 2 are new mathematical results linking HodgeRank and Perron Rank, two popular methods used in pairwise ranking. This converts a large class of the score recovery problem into a parametric optimization question with geometric interpretations, contributing towards the solving of specific instances of this problem.

1. PROOFS

1.1. Proof of Theorem 1. Equip \mathbb{R}^n with the Euclidean norm. Let $\|\cdot\|$ denote the operator norm if the argument is a matrix, and the ℓ_2 norm if the argument is a vector. Let \mathbf{I} be the $n \times n$ identity matrix. Our proof of Theorem 1 relies on the following proposition, which gives a linear approximation to the ‘error term’ of the principal eigenvector when X is not far from being a strongly transitive matrix.

Proposition 3. Fix a vector $s \in \mathbb{R}_+^p$. Let $\xi := [\xi_{ij}] := [s_j/s_i \cdot X_{ij}]$. Suppose there exists a constant $\kappa \geq 1$ such that the $n \times n$ matrix $\Xi = \xi - \kappa \mathbf{1}\mathbf{1}^T - (1 - \kappa)\mathbf{I}$ satisfies $\rho := (\frac{2\|\Xi\|}{n\kappa - 2\|\Xi\|})^2 < \frac{1}{2}$. Then

$$v(X) = s \cdot (\mathbf{1} + \frac{1}{\kappa n}(r - \bar{r}\mathbf{1}) + \epsilon) \quad (1)$$

where $r = (\Xi \cdot \mathbf{1})$ is the row sum of Ξ , $\bar{r} = \frac{1}{n} \sum_{j=1}^n r_j$ is its mean, and $\|\epsilon\| < \frac{\rho}{1-\rho} \cdot \frac{\|\Xi\|}{\kappa\sqrt{n}}$.

This proposition is interesting in itself. One could think of s as the true score vector, $[s_i/s_j]$ as the true multiplicative comparison matrix, and ξ as the multiplicative perturbation with centered version Ξ , so defined since $v(\kappa \mathbf{1}\mathbf{1}^T + (1 - \kappa)\mathbf{I}) = \mathbf{1}$ for all $\kappa \geq 1$. If $\Xi \equiv 0$ then the perturbation does not affect the principal eigenvector, and $v(X)$ is exactly s . The proposition states that if the centered perturbation Ξ is small, as measured by ρ and $\frac{\|\Xi\|}{\kappa\sqrt{n}}$, then $v(X)$ differs from the true multiplicative score s only by a linear factor plus a lower order term. Unfortunately this proposition cannot be applied easily in practice since it requires knowledge of κ , which depends on the choice of s and in general neither are easy to find. For the proof of Theorem 1 it is sufficient to choose $\kappa = 1$.

Proof of Proposition 3. By [19, Lemma 2.2], $v(X) = s \cdot v(\xi)$, hence one can assume without loss of generality that $s \equiv \mathbf{1}$. Write $\xi = \kappa \mathbf{1}\mathbf{1}^T + (\Xi + (1 - \kappa)\mathbf{I})$. Then $\kappa \mathbf{1}\mathbf{1}^T$ is a rank one matrix with one non-zero eigenvalue $p\kappa$, corresponding to the normalized eigenvector

$\frac{1}{\sqrt{n}}\mathbf{1}$. Let $Y \in \mathbb{R}^{n \times (n-1)}$ be an orthonormal basis of the zero eigenspace of $\kappa\mathbf{1}\mathbf{1}^T$. From standard results in perturbed linear operators (see, for example, [18, Theorem 2.7]), $\rho < \frac{1}{2}$ implies $v(\xi) = \frac{1}{\sqrt{n}}\mathbf{1} + YP + \frac{\epsilon}{\sqrt{n}}$, where $P := \frac{1}{n\sqrt{n\kappa}}Y^T(\Xi + (1 - \kappa)\mathbf{I})\mathbf{1}$, and the error term is bounded by

$$\|\epsilon\| \leq \frac{\rho}{1 - \rho} \cdot \frac{\|\Xi\|}{n\kappa}.$$

Since $YY^T = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T$, and since $v(\xi)$ is only defined up a multiplicative constant, we have

$$\begin{aligned} v(\xi) &= \frac{1}{\sqrt{n}} \left(\mathbf{1} + \frac{1}{n\kappa} (\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T) (\Xi + (1 - \kappa)\mathbf{I})\mathbf{1} + \epsilon \right) \\ &= \mathbf{1} + \frac{1}{n\kappa} (\Xi\mathbf{1} - \frac{1}{n}\mathbf{1}\mathbf{1}^T\Xi\mathbf{1}) + \epsilon = \mathbf{1} + \frac{1}{n\kappa} (r - \bar{r}) + \epsilon \end{aligned}$$

□

Proof of Theorem 1. Under the notation of Proposition 3, let $\kappa = 1$, $s \equiv \mathbf{1}$, so $X = \xi$. Define $A = [A_{ij}] = [\log X_{ij}]$. Then

$$\Xi_{ij}^{(k)} = \exp(kA_{ij}) - 1 = k \cdot \left(A_{ij} + \sum_{t \geq 2} \frac{k^{t-1} A_{ij}^t}{t!} \right).$$

Hence for k close to 0, $\|\Xi^{(k)}\| = k\|A\| + o(k)$ and $\frac{\rho}{1-\rho} = O(k^2)$, and Proposition 3 applies. Note that $r = \Xi^{(k)}\mathbf{1} = k \cdot A\mathbf{1} + O(k^2)$, $\bar{r}\mathbf{1} = k \cdot \mathbf{1}\mathbf{1}^T A\mathbf{1} + O(k^2)$, and $\epsilon = O(k^3)$. After applying the logarithm to Equation 1, Taylor expansion gives

$$\lim_{k \rightarrow 0} \frac{1}{k} \log v(X^k) = \lim_{k \rightarrow 0} \left(\frac{1}{kn} r - \frac{1}{kn} \bar{r} + O(k^2) \right) = \frac{1}{n} A\mathbf{1} - \frac{1}{n} \mathbf{1}\mathbf{1}^T A\mathbf{1} = \frac{1}{n} \log h(X)$$

since the HodgeRank vector is defined only up to additive constants. □

1.2. Proof of Theorem 2. It is sufficient to prove equivalent statements on \mathcal{K} for the map V_k . The fibers of this map are the set of positive matrices whose principal eigenvector is a fixed vector $w \in \mathbb{R}_+^n$, that is

$$\mathcal{K}(w) := \bigsqcup_{\mu \in (0, \infty)} \mathcal{K}(w, \mu) := \bigsqcup_{\mu \in (0, \infty)} \{X \in \mathcal{K} : v(X) = w, \lambda(X) = \mu\}$$

where \bigsqcup denote disjoint set union. We call this the *one-dimensional real positive Kalman variety*, motivated by the definition of the Kalman variety in Ottaviani and Sturmfels [14]. In general real Kalman varieties are difficult to characterize, however the one-dimensional positive variety admits a simple description.

Corollary 4. *For each fixed pair (w, λ) , let $\Psi_{w, \lambda} : \mathcal{K} \rightarrow \mathcal{K}$ be the map $[X_{ij}] \mapsto [Y_{ij}] := [\frac{X_{ij}w_j}{\lambda w_i}]$. Then*

$$\Psi_{w, \lambda}(\mathcal{K}(w, \lambda)) = \bigoplus_{i=1}^n (\Delta_{n-1})_i$$

where each $(\Delta_{n-1})_i$ is the interior of the $(n-1)$ -dimensional simplex on the $n \cdot (i-1) + 1$ to $n \cdot i$ coordinate of $\mathbb{R}^{n \times n}$.

Proof of Theorem 2. The first statement follows by [19, Lemma 2.2] (which is a direct computation). By Corollary 4, $V_k^{-1}(\mathbf{1}) = \mathbb{R} \cdot \bigoplus_{i=1}^n (\Delta_{n-1})_i$. By taking log, one obtains the stated result for $\tilde{V}_k^{-1}(\mathbf{0})$. The case for $k = \infty$ follows from the convergence of the real amoebas to the zero set of the tropical eigenvectors map. As k approaches 0, each component $S_i(k)$ is flattened, and a little calculation shows that the each component converges, up to a translation by a large constant times the $(1, \dots, 1)$ vector, to its tangent plane at $a_{i1} = \dots = a_{in}$, which has equation $\sum_{j=1}^n a_{ij} = 0$. This proves the theorem. \square

REFERENCES

- [1] M. Akian, R. Bapat, and S. Gaubert. Asymptotics of the Perron eigenvalue and eigenvector using max-algebra. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 327(11):927 – 932, 1998.
- [2] G. B. Crawford. The geometric mean procedure for estimating the scale of a judgement matrix. *Mathematical Modelling*, 9:327–334, 1987.
- [3] Geir Dahl. A method for approximating symmetrically reciprocal matrices by transitive matrices. *Linear Algebra Appl.*, 403:207–215, 2005.
- [4] Y. Dong, G. Zhang, W. Hong, and Y. Xu. Consensus models for AHP group decision making under row geometric mean prioritization method. *Decision Support Systems*, 49:281 – 289, 2010.
- [5] John Duchi, Lester Mackey, and Michael Jordan. On the consistency of ranking algorithms. Technical Report UCB/EECS-2010-56, EECS Department, University of California, Berkeley, May 2010.
- [6] L. Elsner and P. van den Driessche. Max-algebra and pairwise comparison matrices. *Linear Algebra and its Applications*, 385:47 – 62, 2004.
- [7] L. Elsner and P. van den Driessche. Max-algebra and pairwise comparison matrices, ii. *Linear Algebra and its Applications*, 432(4):927 – 935, 2010.
- [8] A. Farkas. The analysis of the principal eigenvector of pairwise comparison matrices. *Acta Polytechnica Hungarica*, 4:99–115, 2007.
- [9] J. Fürnkranz and E. Hüllermeier. *Preference Learning*. Springer, 2010.
- [10] Dorit S. Hochbaum. Ranking sports teams and the inverse equal paths problem, 2006.
- [11] Eyke Hüllermeier and Johannes Fürnkranz. Learning label preferences: Ranking error versus position error. In *IN PROC. IDA 2005*, pages 180–191. Springer-Verlag, 2005.
- [12] X. Jiang, L-H. Lim, Y. Yao, and Y. Ye. Statistical ranking and combinatorial hodge theory. *Mathematical Programming*, 127:203–244, 2011. 10.1007/s10107-010-0419-x.
- [13] D. Maclagan and B. Sturmfels. Introduction to Tropical Geometry. Manuscript, 2009.
- [14] G. Ottaviani and B. Sturmfels. Matrices with eigenvectors in a given subspace. arXiv:1012.1016v1, 2010.
- [15] D. A. Saari. A new way to analyze paired comparison rules. Manuscript, 2010.
- [16] T. L. Saaty. *The Analytic Hierarchy Process*. McGrawHill, New York, 1980.
- [17] Thomas L. Saaty. Rank according to perron: A new insight. *Mathematics Magazine*, 60(4):pp. 211–213, 1987.
- [18] G. W. Stewart and J. Sun. *Matrix perturbation theory*. Academic Press, Inc, 1990.
- [19] N. M. Tran. Pairwise ranking: choice of method can produce arbitrarily different rank order. arXiv:1103.1110, 2011.

DEPARTMENT OF STATISTICS, UC BERKELEY, CA 94720, USA
E-mail address: tran@stat.berkeley.edu